



# インフルエンザB型患者数動態の数理統計的考察

著者	鈴木 北斗
発行年	2019
URL	<a href="http://hdl.handle.net/10236/00028903">http://hdl.handle.net/10236/00028903</a>

## インフルエンザ B 型患者数動態の数理統計的考察

関西学院大学大学院 理工学研究科  
数理科学専攻 昌子研究室 鈴木 北斗.

## 1 研究背景

インフルエンザは季節性の流行を引き起こす感染症である [1]。主としてヒトに感染するインフルエンザは A 型、B 型の 2 種類であるが両者は抗原性が異なるため、一方に感染すれば一方への免疫が得られるといった交差反応性はない。一方で、A 型と B 型に重感染した際はウイルス間の干渉により一方のウイルスの増殖が抑制されることが知られている [3]。また流行開始時期が早いのは A 型であることが多いため、社会動態的観点からも B 型の流行は抑制されるのではないかと考えられる。

インフルエンザの動態解析は古くから行われてきたが、A 型、B 型両方の合計値及び感染力の強い A 型のみに関する分析を考えたものが多い。本研究では、A 型 B 型間に相互作用性があることの統計的な考察を行い、また B 型の予測に A 型の動態も取り入れることで B 型動態予測精度が上げられないかと分析・予測を行った。

## 2 分析手法

分析・予測にあたって国立感染症研究所にて公開されている「インフルエンザウイルス分離・検出速報」の週別データを使用する [1]。気象データは、気象庁にて公開されている「過去の気象データ (東京)」を使用する [2]。

## 2.1 VAR モデル (ベクトル自己回帰モデル)

VAR モデルは AR モデルをベクトルに一般化した多変量確率過程の時系列モデルである [4]。特徴は、複数の変数について各変数間の影響をモデルに組み込むことができる点にある。また、VAR モデルは変数間の相互作用関係の分析に、グレンジャー因果性検定やインパルス応答関数などの強力なツールを提供する。

## 2.2 グレンジャー因果性検定

グレンジャー因果性とはデータだけから因果性を判断する概念である。具体的には、「現在と過去の  $x$  の値だけに基いた将来の  $x$  の予測」と「現在と過去の  $x$  と  $y$  の値に基いた将来の  $x$  の予測」を比較し、後者の平均二乗誤差の方が小さくなる場合、 $y_t$  から  $x_t$  へグレンジャー因果性が存在すると言われる [4]。

## 2.3 インパルス応答関数

インパルス応答関数とは、ある変数に対するショックが、その変数や他の変数に与える影響を定量的に分析するためのツールである [4]。VAR モデルではショックの識別方法によって複数のインパルス応答関数が存在し、今回用いるのは直交化インパルス応答関数である。

## 2.4 分散分解

インパルス応答関数のように変数間の動学的関係を定量的に解析する手法の一つとして、分散分解が挙げられる [4]。これにより、各変数が互いにどの程度の影響を及ぼしあっているのかを定量的に解析することができる。

## 3 分析結果

A 型、B 型に関する VAR モデル推定の結果、BIC 基準で次数は 7 となった。以下これにより分析を行う。

## 3.1 グレンジャー因果性検定の結果

因果性検定の結果、「A 型から B 型」にのみ有意水準 1 % で因果性が存在するとなった。これにより B 型の予測に A 型を用いるための根拠が得られた。

## 3.2 インパルス応答関数及び分散分解の結果

インパルス応答関数の計算結果を図 1、図 2 で示す。図 1 より A 型の B 型に対する感染拡大の促進効果を読み取れ、これは生物学的に抑制効果のある両者であるが、統計的には互いに促進効果をもつことを示している。

また、分散分解の結果を図 3 に示す。これは上から「気温、絶対湿度、A 型報告数、B 型報告数」に対する結果であり、分析には A 型、B 型に相関分析にて相関係数が高かった「絶対湿度」「気温」を含めた 4 変数に関して推定した次数 6 の VAR モデルを用いた。これにより、B 型の動態は B 型、A 型要因が大きいことが分かる。

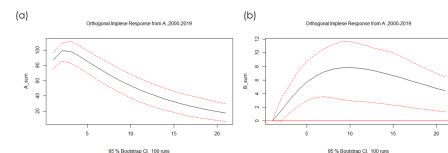


図 1 インパルス応答;(a)A 型から A 型, (b)A 型から B 型

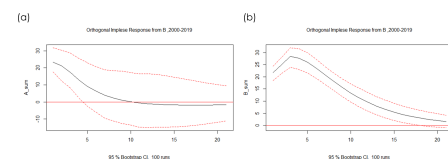


図 2 インパルス応答;(a)B 型から A 型, (b)B 型から B 型

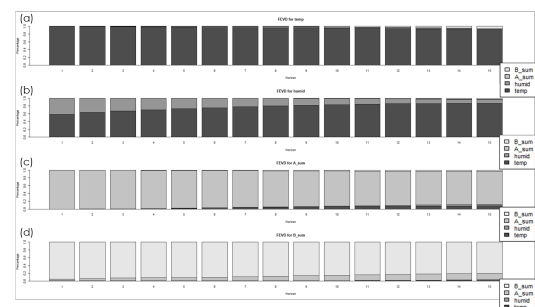


図 3 分散分解の結果; (a) 気温, (b) 絶対湿度, (c)A 型報告数, (d)B 型報告数

## 4 新たな予測アルゴリズムの提案

前節で示された統計的指標を用いて、B 型の 1 か月程度の動態予測をする新たなアルゴリズムを提案する。

### 4.1 インパルス応答関数の影響パラメータ推定

A 型の患者数が与えられたとき、前節で計算したインパルス応答により、B 型の患者数動態が予測できるだろうと考え、後述のような B 型時系列分布モデルを考える。図 1 右側の (b) の計算結果を数値で捉え、これらを順番に  $i_1, i_2, \dots$  と置く。また、時刻  $t$  における B 型報告数を  $y_{B,t}$ 、A 型報告数を  $y_{A,t}$ 、 $a, b, s$  をパラメータとすると、 $y_{B,t}$  のモデル式を以下で定める。

$$y_{B,t} = s \sum_{k=1}^p i_k y_{A,t-k} \frac{1}{1 + e^{a(k-b)}} \quad (1)$$

パラメータ  $a, b, s$  は、推定結果と報告数の間における RMSE を最小にするものを推定する。 $p$  はインパルス応答をどの時点まで用いるかを表し、今回は各インフルエンザの流行期間を基準とし、3 か月程度の  $p = 14$  とした。また、実際の B 型報告数を  $Y_{B,t}$  とし、各誤差に対する重みづけを  $\sqrt{Y_{B,t}}/10$  で定式化した。以上の仮定の下で、このモデルは A 型から B 型に対するインパルス応答の値を B 型の報告数に回帰させる形式をとる。

### 4.2 推定結果

データを 2009 年を境界として前半 (2000-2008 年)、後半 (2011-2018 年) の 2 つに分けた。そして「(1) 前半をトレーニングデータ、後半をテストデータとした場合」「(2) 後半をトレーニングデータ、前半をテストデータとした場合」の 2 つに場合を分け、それぞれに対してモデルを推定した。また、パラメータ推定には Excel のソルバーを用いた。(1) について図 4, 図 5 に結果を示す。一部規模の大きさを捉えられていない箇所もあるが、ピークの時期は大方捉えられていることが確認できる。

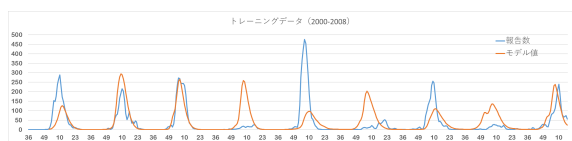


図 4 トレーニングデータに対するフィッティング

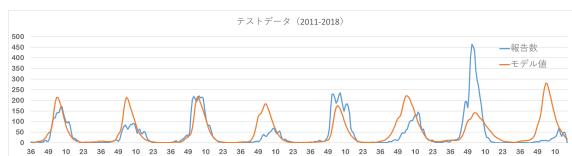


図 5 テストデータに対するフィッティング

### 4.3 SIR モデルベース提案アルゴリズム

上述のモデルと、感染動態モデルとして有名な SIR モデルを用いて、先 1 ヶ月程度の患者数予測アルゴリズムを作成した [5]。まず、B 型動態予測には A 型報告数の予測も必要になるため、ここでは SIR モデルに基づき A 型を予測した。

## 5 結果

予測結果の一部を図 6 に示す。SIR モデルに基づく式を用いて 3 週先までの A 型報告数を予測し、それを用いて 4 週先の B 型報告数を予測した。2000 年度の B 型患者数の急激な増加には対応できにくいものとなっているが、その他の動態の傾向予測は上手く行えているようだ。

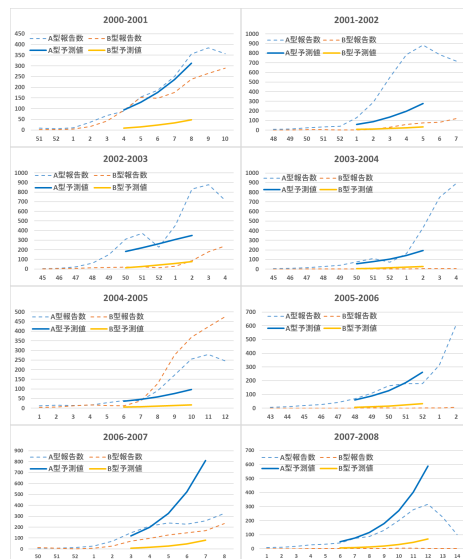


図 6 後半をトレーニングデータにした際の推定結果

## 6 議論

本研究では、感染症の 1 つであるインフルエンザの感染者数動態を数理的・統計的視点により解析を行った。そして、得られた統計的結果を組み込んだ数理モデルを作成し、患者数予測アルゴリズムを作成、検証した。

統計解析により、生物学的事実と反する A 型と B 型間の数値的な促進作用の存在を示した。同時に、A 型から B 型へのグレンジャー因果性の存在も示した。また、より具体的なアプリケーションとして、現場で実用的な患者数動態予測アルゴリズムを提案した。これは得られた統計的指標、及び SIR モデルをベースとした予測ツールである。統計の各手法では、各年毎のデータセットを用いての推論になるが、ここで開発した手法はシーズン途中での先 1 ヶ月の患者数予測を行うものとなっている。

しかし、提案手法では A 型の予測も必要であり、今後は A 型の予測手法も確立していく必要がある。とはいえ、インフルエンザは A 型が先に流行する傾向があるため、B 型が流行を始める前に B 型の動態予測を行うことができる点は、提案手法の優れている点だと考える。

## 参考文献

- [1] 国立感染症研究所 ホームページ
- [2] 気象庁, 過去の気象データ・ダウンロード
- [3] 正木明夫, 松浦久美子, 香取幸治, 永井美之: AH 3 型と B 型インフルエンザウイルスによる混合感染の 1 例について. 小児感染免疫, 15 (2003), 199-205.
- [4] 沖本竜義, 「経済・ファイナンスデータの計量時系列分析」 (2010), 朝倉書店.
- [5] Murray J.D., *Mathematical Biology*, 2002, Springer.